Main Subject Detection via Adaptive Feature Selection

Cuong Vu and Damon Chandler

Image Coding and Analysis Lab
Oklahoma State University
Main Subject Detection is easy for human ...
Outline

• Introduction
• Algorithm
• Results and Analysis
• Conclusion and Future work
Outline

• Introduction
• **Algorithm**
• Results and Analysis
• Conclusion and Future work
Outline

• Introduction
• Algorithm
• Results and Analysis
• Conclusion and Future work
Outline

• Introduction
• Algorithm
• Results and Analysis
• Conclusion and Future work
Outline

• Introduction
• Algorithm
• Results and Analysis
• Conclusion and Future work
Introduction

• Photographers convey their ideas in photos via one or more main subjects.
• Human visual system pays more attention to some parts of images.
• Applications of Main Subject Detection (MSD)
  – Auto cropping a photo
  – Image compression
  – Unequal error protection
  – Object recognition
  – Etc.
Previous work

• Algorithms to detect the main subject in an image
Previous work

- Algorithms predict visual fixation
Outline

• Introduction
• Algorithm
• Results and Analysis
• Conclusion and Future work
Algorithm

Adaptive Feature Selection & Weighted Sum

Multi stage Refinement

Feature Measurement

Saliency Map & Bounding Box
Outline

• Introduction

• Algorithm
  o Feature Measurement

• Results and Analysis

• Conclusion and Future work
Features

- Sharpness
- Color Distance
- Lightness Distance
- Contrast
- Edge Strength
Sharpness\textsuperscript{[1]}

Original image

Convert to gray scale image

32×32 block
24 pixels overlap

Spectral
sharpness map $S_1$

8×8 block
No overlap

Spatial sharpness
map $S_2$

Final sharpness
map $S_3$

Geometric
mean

Color Distance

(1) Original image

(2) Convert to $a^*$, $b^*$

Measure dist. from avg. block $a^*$, $b^*$ to avg. bkgnd $a^*$, $b^*$

(3) 8x8 blocks w/ 50% overlap

Local color dist. map
Lightness Distance

(1) Original image

(2) Convert to $L^*$

(3) 8x8 blocks w/ 50% overlap

Measure dist. from avg. block $L^*$, to avg. bkgnd $L^*$

Local lightness dist. map
RMS Contrast

(1) Original image

(2) Convert to luminance

(3) 8x8 blocks w/ 50% overlap

Measure avg. block RMS contrast

Local contrast map
Edge Strength

(1) Original image

(2) Run Robert’s edge detector

(3) 8x8 blocks w/ 50% overlap

Measure avg. block edge pixels

Local edge strength map
Outline

• Introduction

• Algorithm
  o Feature Measurement
  o **Adaptive Feature Selection**

• Results and Analysis

• Conclusion and Future work
Adaptive feature selection

Sharpness
Contrast
Lightness
Color
Edge
Adaptive feature selection

Sharpness
Contrast
Lightness
Color
Edge
Adaptive feature selection

How to measure the usefulness of a feature?

- Sharpness
- Contrast
- Lightness

- Color
- Edge
Cluster distance

Clustered

Non-clustered

The more clustered, the better!!!
Centroid

A saliency pixel

Saliency region, includes \( P \) pixels

Distance \( d_j \)

Centroid

Cluster Distance

\[
\beta = \frac{\sum d_j}{P^\alpha}
\]
Cluster distance

Low cluster distance

Cluster Distance = 11.07

High cluster distance

Cluster Distance = 18.56
Outline

• Introduction

• Algorithm
  o Feature
  o Adaptive Feature Selection
    o Multi-stage Refinement

• Results and Analysis

• Conclusion and Future work
Stage 1: Measure Features

Spatial weighting

Sharpness

Lightness Distance

Contrast

Color Distance

Edge Strength
Stage 1: Measure Features

\[ w_i = \begin{cases} 
1, & \text{if } \beta_i = \tilde{\beta}_5 \\
2/3, & \text{if } \beta_i = \tilde{\beta}_4 \\
1/3, & \text{if } \beta_i = \tilde{\beta}_3 \\
0, & \text{otherwise} 
\end{cases} \]

where \( \tilde{\beta} = \text{sort } \{\beta_i\} \)
✓ Saliency pixels: Greater than $T = 1.5 \times \text{mean}(map)$
Stage 2: Refine Features

- Sharpness
- Lightness Distance
  - To the new background
- Contrast
- Color Distance
  - To the new background
- Edge Strength

Background

Spatial weighting
Stage 2: Refine Features

\[ w_i = \begin{cases} 
1, & \text{if } \beta_i = \tilde{\beta}_5 \\
\frac{\tilde{\beta}_1 - \tilde{\beta}_4}{\tilde{\beta}_1 - \tilde{\beta}_5}, & \text{if } \beta_i = \tilde{\beta}_4 \\
0, & \text{otherwise}
\end{cases} \]

where \( \tilde{\beta} = \text{sort}\{\beta_i\} \)
Stage 2 map

Stage 2 rectangle

✓ Saliency pixels: Greater than $T = 2 \times \text{mean}(\text{map})$
✓ Take only 75% of saliency pixels
Stage 3: Refine Features Again

Spatial weighting

Foreground

Sharpness

Lightness Distance
To the new foreground

Contrast

Color Distance
To the new foreground

Edge Strength
Stage 3: Refine Features Again (Distance to the foreground)

\[ w_i = \begin{cases} 
1, & \text{if } \beta_i = \tilde{\beta}_5 \\
\frac{\tilde{\beta}_1 - \tilde{\beta}_4}{\beta_1 - \tilde{\beta}_5}, & \text{if } \beta_i = \tilde{\beta}_4, \text{ where } \tilde{\beta} = \text{sort } \{\beta_i\} \\
0, & \text{otherwise}
\end{cases} \]
✓ Saliency pixels: Greater than $T = 1.5 \times \text{mean}(\text{map})$
Outline

• Introduction
• Algorithm
• Results and Analysis
• Conclusion and Future work
MSD Results on Objects
MSD Results on Plants
MSD Results on People
MSD Results on Animals
MSD Failure Cases
Image testing sets

- MSRA database\(^2\) : 5000 images (set B) with bounding box groundtruth

- EPFL database\(^3\) : 1000 images from MSRA database with object-contour based groundtruth\(^3\)

---


\(^3\) R. Achanta et al., “Frequency-turned Salient Region Detection”, IEEE Conference on Computer Vision and Pattern Recognition, 2009
Evaluation Criteria

• Boundary-based Criterion:
  ➢ BDE: Boundary Displacement Error\(^4\) between the boundaries of Detected and Groundtruth rectangles

• Region-based Criteria
  ➢ Precision = \( \frac{A(D \cap G)}{A(D)} \)
  ➢ Recall = \( \frac{A(D \cap G)}{A(G)} \)
  ➢ F-measure = \( \frac{(1 + \alpha) \times P \times R}{\alpha \times P + R} \)

Overall Results – MSRA database

![Bar chart showing BDE values for different methods]

Better


Overall Results – MSRA database

Better


Precision  Recall  F-Measure
Overall Results – EPFL database

![Graph showing overall results for different methods]

- **Y. Ma et al. [5]**
- **L. Itti et al. [6]**
- **J. Harel et al. [7]**
- **X. Hou et al. [8]**
- **R. Achanta et al. [3]**
- **AFS**
- **AFS_MS**

For references:


## Average weighs of features

<table>
<thead>
<tr>
<th></th>
<th>Sharpness</th>
<th>Lightness Distance</th>
<th>Contrast</th>
<th>Color Distance</th>
<th>Edge Strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 1</td>
<td>0.23</td>
<td>0.22</td>
<td>0.10</td>
<td>0.24</td>
<td>0.21</td>
</tr>
<tr>
<td>Stage 2</td>
<td>0.32</td>
<td>0.23</td>
<td>0.06</td>
<td>0.32</td>
<td>0.07</td>
</tr>
<tr>
<td>Stage 3</td>
<td>0.32</td>
<td>0.21</td>
<td>0.08</td>
<td>0.32</td>
<td>0.07</td>
</tr>
<tr>
<td>Overall</td>
<td>0.29</td>
<td>0.22</td>
<td>0.08</td>
<td>0.29</td>
<td>0.12</td>
</tr>
</tbody>
</table>
Conclusion and Future work

• Relatively simple low-level features can be effective for MSD if they are combined in an adaptive and multi-stage fashion.
• Our results are very promising.
• This adaptive feature selection can be a useful strategy for other applications.
• Future works:
  – Improve the algorithm: new feature, using segmentation, etc
  – Multiple MSD
Thank you!

http://vision.okstate.edu
{cuong.vu, damon.chandler}@okstate.edu
Stage 1: Measure Features

Spatial weighting

Sharpness
$\omega = 1$

Lightness Distance
$\omega = 1/3$

Contrast

Color Distance

Edge Strength
$\omega = 2/3$
Stage 2: Refine Features

Background

Spatial weighting

Sharpness
$\omega = 0.826$

Lightness Distance

Contrast

Color Distance

Edge Strength
$\omega = 1$
Stage 3: Refine Features Again

Foreground

Spatial weighting

Sharpness
\( \omega = 0.95 \)

Lightness Distance

Contrast
\( \omega = 1 \)

Color Distance

Edge Strength